

Contents

1. Relative Information – What For?	1
1.1 Information Theory, What Is It?	1
1.1.1 Summary of the Story	1
1.1.2 Communication and Information	2
1.2 Information and Natural Language	3
1.2.1 Syntax, Semantics, Lexeme	3
1.2.2 Information, Learning, Dislearning	3
1.3 Prerequisites for a Theory of Information	4
1.3.1 Relativity of Information	4
1.3.2 Negative Information	5
1.3.3 Entropy of Form and Pattern	5
1.3.4 Information and Thermodynamics	6
1.3.5 Information and Subjectivity	6
1.4 Information and Systems	7
1.4.1 A Model of General Systems	8
1.4.2 A Model of Relative Information	8
1.4.3 A Few Comments	9
1.5 How We Shall Proceed	10
1.5.1 Aim of the Book	10
1.5.2 Subjective Information and Relative Information	10
1.5.3 Minkowskian Observation of Events	10
1.5.4 A Unified Approach to Discrete Entropy and Continuous Entropy	11
1.5.5 A Word of Caution to the Reader	11
 2. Information Theory – The State of the Art	 12
2.1 Introduction	12
2.2 Shannon Measure of Uncertainty	13
2.2.1 The Probabilistic Framework	13
2.2.2 Shannon Informational Entropy	13
2.2.3 Entropy of Random Variables	15
2.3 An Intuitive Approach to Entropy	15
2.3.1 Uniform Random Experiments	15
2.3.2 Non Uniform Random Experiments	16
2.4 Conditional Entropy	17
2.4.1 Framework of Random Experiments	17
2.4.2 Application to Random Variables	18
2.5 A Few Properties of Discrete Entropy	19

2.6	Prior Characterization of Discrete Entropy	20
2.6.1	Properties of Uncertainty	21
2.6.2	Some Consequences of These Properties	22
2.7	The Concept of Information.....	22
2.7.1	Shannon Information	22
2.7.2	Some Properties of Transinformation	23
2.7.3	Transinformation of Random Variables	25
2.7.4	Remarks on the Notation	25
2.8	Conditional Transinformation.....	25
2.8.1	Main Definition.....	25
2.8.2	Some Properties of Conditional Transinformation	26
2.8.3	Conditional Transinformation of Random Variables	27
2.9	Renyi Entropy	27
2.9.1	Definition of Renyi Entropy	27
2.9.2	Meaning of the Renyi Entropy	28
2.9.3	Some Properties of the Renyi Entropy	29
2.10	Cross-Entropy or Relative Entropy	29
2.10.1	The Main Definition.....	29
2.10.2	A Few Comments	30
2.11	Further Measures of Uncertainty	31
2.11.1	Entropy of Degree c	31
2.11.2	Quadratic Entropy	31
2.11.3	R-norm Entropy	32
2.11.4	Effective Entropy	32
2.12	Entropies of Continuous Variables	33
2.12.1	Continuous Shannon Entropy.....	33
2.12.2	Some Properties of Continuous Entropy	34
2.12.3	Continuous Transinformation	35
2.12.4	Further Extensions	36
2.13	Hatori's Derivation of Continuous Entropy.....	37
2.14	Information Without Probability	39
2.14.1	A Functional Approach	39
2.14.2	Relative Information.....	40
2.15	Information and Possibility.....	41
2.15.1	A Few Prerequisites.....	41
2.15.2	A Measure of Uncertainty Without Probability	42
2.16	Conclusions	42
3.	A Critical Review of Shannon Information Theory.....	44
3.1	Introduction	44
3.2	On the Invariance of Measures of Information.....	45
3.3	On the Modelling of Negative Transinformation.....	46
3.3.1	Classification of Terms.....	46
3.3.2	The Problem of Modelling "True" and "False"	46
3.3.3	Error-Detecting Codes	47
3.4	On the Symmetry of Transinformation	49
3.4.1	A Diverting Example	49

3.4.2	Application of Information Theory	49
3.4.3	On the Symmetry of Transinformation	50
3.4.4	On a Possible Application of Renyi Entropy	51
3.5	Entropy and the Central Limit Theorem	51
3.5.1	The Central Limit Theorem	52
3.5.2	An Information-Theoretical Approach to the Central Limit Theorem.....	53
3.5.3	Relation with Thermodynamics	54
3.5.4	Continuous Entropy Versus Discrete Entropy	54
3.6	On the Entropy of Continuous Variables	55
3.6.1	The Sign of the Continuous Entropy	55
3.6.2	A Nice Property of Continuous Entropy	56
3.7	Arguments to Support Continuous Entropy.....	56
3.7.1	On the Negativeness of Continuous Entropy.....	56
3.7.2	On the Non-invariance of Continuous Entropy.....	56
3.7.3	Channel Capacity in the Presence of Noise	57
3.8	The Maximum Entropy Principle.....	58
3.8.1	Statement of the Principle.....	58
3.8.2	Some Examples	59
3.9	Arguments to Support the Maximum Entropy Principle	60
3.9.1	Information-Theoretical Considerations	60
3.9.2	Thermodynamic Considerations	61
3.9.3	Axiomatic Derivation	61
3.9.4	A Few Comments	61
3.10	Information, Syntax, Semantics	62
3.10.1	On the Absolute Nature of Information.....	62
3.10.2	Information and Natural Language	63
3.11	Information and Thermodynamics.....	63
3.11.1	Informational and Thermodynamic Entropy	63
3.11.2	Thermodynamic Entropy of Open Systems	64
3.12	Conclusions	65
4.	A Theory of Relative Information	66
4.1	Introduction	66
4.2	Observation, Aggregation, Invariance	67
4.2.1	Principle of Aggregation	67
4.2.2	Principle of Invariance.....	68
4.2.3	A Few Comments	68
4.3	Observation with Informational Invariance	69
4.4	Euclidean Invariance.....	70
4.4.1	Orthogonal Transformation.....	70
4.4.2	Application to the Observation of Probabilities	71
4.4.3	Application to the Observation of Classes	72
4.5	Minkowskian Invariance.....	74
4.5.1	Lorentz Transformation	74
4.5.2	Application to the Observation of Probabilities	75
4.5.3	Application to the Observation of Classes	75

4.6	Euclidean or Minkowskian Observation?	76
4.6.1	Selection of the Observation Mode	76
4.6.2	Application to the [Uncertainty, Information] Pair	77
4.7	Information Processes and Natural Languages	78
4.7.1	On the Absoluteness of Information	78
4.7.2	The Information Process	78
4.7.3	Natural Language	79
4.7.4	Information and Natural Language	80
4.8	Relative Informational Entropy	80
4.8.1	Introduction to the Relative Observation	80
4.8.2	Informational Invariance of the Observation	81
4.8.3	Relative Entropy	82
4.8.4	Comments and Remarks	82
4.9	Conditional Relative Entropy	84
4.9.1	Relative Entropy of Product of Messages	84
4.9.2	Composition Law for Cascaded Observers	85
4.9.3	Relative Entropy Conditional to a Given Experiment	86
4.9.4	Applications to Determinacy	87
4.9.5	Comparison of $H(\beta/\alpha)$ with $H_r(\beta/\alpha)$	87
4.10	On the Meaning and the Estimation of the Observation Parameter	87
4.10.1	Estimation of the Observation Parameter	88
4.10.2	Practical Meaning of the Observation Parameter	89
4.10.3	On the Value of the Observation Parameter $u(R)$	90
4.11	Relative Transinformation	90
4.11.1	Derivation of Relative Transinformation	90
4.11.2	Some Properties of the Relative Transinformation	91
4.11.3	Relative Entropy and Information Balance	92
4.11.4	Application to the Encoding Problem	93
4.12	Minkowskian Relative Transinformation	95
4.12.1	Definition of Minkowskian Relative Transinformation	95
4.12.2	Some Properties of Minkowskian Relative Transinformation	96
4.12.3	Identification via Information Balance	96
4.13	Effect of Scaling Factor in an Observation with Informational Invariance	97
4.14	Comparison with Renyi Entropy	98
4.14.1	Renyi Entropy and Relative Entropy	98
4.14.2	Transinformation of Order c and Relative Transinformation	99
5.	A Theory of Subjective Information	100
5.1	Introduction	100
5.2	Subjective Entropy	101
5.2.1	Definition of Subjective Entropy	101
5.2.2	A Few Remarks	102
5.3	Conditional Subjective Entropy	102

5.3.1	Definition of Conditional Subjective Entropy	102
5.3.2	Application to Determinacy	103
5.3.3	A Basic Inequality	104
5.4	Subjective Transinformation	105
5.4.1	Introduction	105
5.4.2	Subjective Transinformation	106
5.4.3	A Few Properties of Subjective Transinformation	107
5.4.4	Application to Independent Random Experiments	109
5.5	Conditional Subjective Transinformation	109
5.5.1	Definition	109
5.5.2	A Few Properties of Subjective Conditional Transinformation	109
5.6	Information Balance	111
5.6.1	Optimum Conditions for Information Balance	111
5.6.2	Non-optimum Conditions for Information Balance	113
5.7	Explicit Expression of Subjective Transinformation	114
5.7.1	Discrete Probability	114
5.7.2	Continuous Probability	115
5.8	The General Coding Problem	115
5.8.1	Preliminary Remarks	115
5.8.2	The General Coding Problem	116
5.8.3	On the Problem of Error Correcting Codes Revisited	118
5.9	Capacity of a Channel	119
5.9.1	The General Model	119
5.9.2	Channel with Noise	120
5.9.3	Channel with Noise and Filtering	121
5.10	Transinformation in the Presence of Fuzziness	123
5.10.1	On the Entropy of a Fuzzy Set	123
5.10.2	Application of Subjective Transinformation	125
5.10.3	The Brillouin Problem	126
5.11	On the Use of Renyi Entropy	126
5.11.1	Renyi Entropy and Subjective Entropy	127
5.11.2	Transinformation of Order c and Shannon Transinformation	127
6.	A Unified Approach to Discrete and Continuous Entropy	128
6.1	Introduction	128
6.2	Intuitive Derivation of "Total Entropy"	129
6.2.1	Preliminary Definitions and Notation	129
6.2.2	Physical Derivation of $H_e(X)$	130
6.3	Mathematical Derivation of Total Entropy	131
6.3.1	The Main Axioms	131
6.3.2	Derivation of Total Entropy	132
6.3.3	On the Expression of the Total Entropy	133
6.4	Alternative Set of Axioms for the Total Entropy	134
6.4.1	Generalization of Shannon Recurrence Equation	134
6.4.2	A Model via Uniform Interval of Definition	134

6.5	Total Entropy with Respect to a Measure	135
6.6	Total Renyi Entropy	135
6.6.1	Preliminary Remarks About Renyi Entropy	135
6.6.2	Axioms for Total Renyi Entropy	136
6.6.3	Total Renyi Entropy	136
6.6.4	Total Renyi Entropy with Respect to a Measure	137
6.7	On the Practical Meaning of Total Entropy	137
6.7.1	General Remarks	137
6.7.2	Total Entropy and Relative Entropy	138
6.7.3	Total Entropy and Subjective Entropy	139
6.8	Further Results on the Total Entropy	139
6.8.1	Some Mathematical Properties of the Total Entropy	139
6.8.2	On the Entropy of Pattern	141
6.9	Transinformation and Total Entropy	142
6.9.1	Total Entropy of a Random Vector	142
6.9.2	Conditional Total Entropy	143
6.9.3	On the Definition of Transinformation	144
6.9.4	Total Kullback Entropy	144
6.10	Relation Between Total Entropy and Continuous Entropy	145
6.10.1	Total Shannon Entropy and Continuous Entropy	145
6.10.2	Total Renyi Entropy and Continuous Renyi Entropy	146
6.10.3	Application to an Extension Principle	147
6.11	Total Entropy and Mixed Theory of Information	147
6.11.1	Background to Effective Entropy and Inset Entropy	148
6.11.2	Inset Entropy is an Effective Entropy	148
7.	A Unified Approach to Informational Entropies	
	via Minkowskian Observation	150
7.1	Introduction	150
7.2	Axioms of the Minkowskian Observation	150
7.2.1	Definition of the Observation Process	151
7.2.2	Statement of the Axioms	151
7.2.3	A Few Comments	152
7.3	Properties of the Minkowskian Observation	153
7.3.1	Learning Process	153
7.3.2	Practical Determination of μ	154
7.3.3	Cascaded Observation of Variables	154
7.4	Minkowskian Observations in $\mathbb{R}^3$	156
7.4.1	Derivation of the Equations	156
7.4.2	A Few Comments	157
7.5	The Statistical Expectation Approach to Entropy	158
7.5.1	Preliminary Remarks	158
7.5.2	Entropy and Statistical Expectation	159
7.6	Relative Entropy and Subjective Entropy	160
7.6.1	Comparison of H_r with H_s	160
7.6.2	Comparison of $T(\beta/\alpha)$ with $J(\beta/\alpha)$	160
7.7	Weighted Relative Entropy	161

7.7.1	Background on Weighted Entropy	161
7.7.2	Weighted Relative Entropy	161
7.7.3	Weighted Relative Entropy of a Vector	163
7.7.4	Total Weighted Relative Entropy	164
7.7.5	Observation and Subjectivity	166
7.8	Weighted Transinformation	166
7.8.1	Some Preliminary Remarks	166
7.8.2	Weighted Relative Conditional Entropy	167
7.8.3	Weighted Transinformation	168
7.8.4	Relation to Renyi Entropy	169
7.9	Weighted Cross-Entropy, Weighted Relative Entropy	171
7.9.1	Background on Kullback Entropy	171
7.9.2	Weighted Kullback Entropy	171
7.10	Weighted Relative Divergence	172
7.11	Application to Continuous Distributions	173
7.11.1	The General Principle of the Derivations	174
7.11.2	Continuous Weighted Relative Entropy	174
7.11.3	Continuous Weighted Kullback Entropy	175
7.11.4	Continuous Weighted Relative Divergence	176
7.12	Transformation of Variables in Weighted Relative Entropy	177
7.13	Conclusions	178
8.	Entropy of Form and Pattern	180
8.1	Introduction	180
8.1.1	Entropy of Form and Fractal Dimension	180
8.1.2	Entropy and Pattern Representation	181
8.1.3	Randomized Definition or Geometrical Approach?	182
8.2	Total Entropy of a Random Vector	182
8.2.1	Definition of the Total Entropy of an m -Vector	182
8.2.2	Some Properties of the Total Entropy of m -Vectors	183
8.3	Total Entropy of a Discrete Quantized Stochastic Process	184
8.3.1	Preliminary Remarks	184
8.3.2	Axioms for the Total Uncertainty of a Discrete Trajectory	185
8.3.3	Expression for the Total Entropy of the Trajectory	186
8.3.4	Some Properties of the Total Entropy of the Trajectory ..	187
8.3.5	A Generalization	187
8.4	Total Renyi Entropy of Discrete Quantized Stochastic Processes ..	188
8.4.1	Total Renyi Entropy of a Vector	188
8.4.2	Derivation of the Total Renyi Entropy	189
8.5	Entropy of Order c Revisited	190
8.5.1	Preliminary Remarks	190
8.5.2	A New Definition of Total Entropy of Order c	191
8.6	Entropy of Continuous White Stochastic Trajectories	191
8.6.1	Notation and Remarks	191
8.6.2	A List of Desiderata for the Entropy of a White Trajectory	192

8.6.3	Towards a Mathematical Expression of the Trajectory Entropy	192
8.6.4	Entropy of Continuous White Trajectories	194
8.7	Entropy of Form and Observation Modes	195
8.7.1	Relative Uncertainty via Observation Modes	195
8.7.2	Classification of Observation Modes	195
8.7.3	Trajectory Entropy via White Observation	196
8.8	Trajectory Entropies of Stochastic Processes	196
8.8.1	Trajectory Shannon Entropy	196
8.8.2	Renyi Entropy of a Stochastic Trajectory	197
8.8.3	A Few Comments	198
8.9	Trajectory Entropy of a Stochastic Process Under Local Markovian Observation	198
8.9.1	Markovian Processes	198
8.9.2	Non-Markovian Processes	200
8.9.3	Transformation of Variables	200
8.9.4	A Few Remarks	202
8.10	Trajectory Entropy of a Stochastic Process Under Global Markovian Observation	203
8.10.1	Markovian Processes	203
8.10.2	Non-Markovian Processes	204
8.10.3	Transformation of Variables	204
8.10.4	A Few Remarks	204
8.11	Trajectory Entropies of Stochastic Vectors	205
8.11.1	Notation and Preliminary Results	205
8.11.2	Trajectory Entropies in $\mathbb{R}^n$	205
8.12	Transinformation of Stochastic Trajectories	206
8.12.1	Definition of Transinformation of Trajectories	206
8.12.2	Transinformation Measures of Stochastic Trajectories	207
8.12.3	Application to the Derivation of Conditional Trajectory Entropies	208
8.13	On the Entropy of Deterministic Pattern	209
8.13.1	Background on Some Results	209
8.13.2	On the Entropy of Deterministic Pattern	210
8.13.3	Dependence of the Entropy upon the Observation Mode	210
8.14	Entropy of a Differentiable Mapping	211
8.14.1	Entropy with Respect to a Family of Distributions	211
8.14.2	Maximum Entropy of a Differentiable Mapping	212
8.14.3	Entropy of a Mapping Defined on a Finite Interval	213
8.14.4	Entropy of a Mapping with an Incomplete Probability Distribution	213
8.15	Entropy of Degree d of Differential Mappings	214
8.15.1	Trajectory Entropy of Degree d	214
8.15.2	Some Properties of Trajectory Entropy of Degree d	215
8.15.3	Practical Meaning of Trajectory Entropy of Degree d	216
8.16	Transinformation Between Differentiable Mappings	217

8.16.1	The Trajectory Entropy of Compositions of Functions . . .	217
8.16.2	Application to Transinformation Between Mappings	218
8.17	Trajectory Entropy of Degree d in Intrinsic Coordinates	218
8.18	Trajectory Entropy of $\mathbb{R}^n \rightarrow \mathbb{R}^n$ Mapping	219
8.19	Trajectory Entropy of Degree d of a Discrete Mapping	219
8.20	Trajectory Entropy and Liapunov Exponent	220
9.	A Theory of Relative Statistical Decision	221
9.1	Introduction	221
9.2	Background on Observation with Information Invariance	222
9.2.1	Euclidean or Minkowskian Invariance?	222
9.2.2	On the Practical Meaning of the Model	223
9.2.3	Syntax and Semantics	224
9.3	Relative Probability	224
9.3.1	Definition of Relative Probability	224
9.3.2	On the Practical Meaning of Relative Probability	226
9.3.3	Semantic Entropy of a Real-Valued Variable	226
9.3.4	Relative Probability of Deterministic Events	227
9.3.5	Relative Probability of the Impossible Event	227
9.4	Composition Laws for Relative Probability	227
9.4.1	Relative Probability of the Complement of an Event	227
9.4.2	Relative Probability of a Deterministic Event	228
9.4.3	Relative Probability of Intersection of Events	229
9.4.4	Relative Probability of the Union of Events	231
9.5	Generalized Maximum Likelihood Criterion	233
9.5.1	The Problem	233
9.5.2	Generalized Maximum Likelihood Criterion	233
9.5.3	Practical Meaning of the Generalized Maximum Likelihood Criterion	233
9.5.4	A Simplified Criterion	234
9.5.5	Another Simplified Decision Criterion	235
9.5.6	Comparison with Another Criterion	236
9.6	Generalized Bayesian Criterion	236
9.6.1	Pattern Recognition	236
9.6.2	Formulation of the Statistical Decision Problem	237
9.6.3	Sharp Decision Rule	238
9.6.4	Soft Decision Rule	238
9.7	Generalized Bayesian Decision and Path Integrals	239
9.7.1	Formulation of the Problem	239
9.7.2	The Main Assumptions and Solution	239
9.8	Error Probability and Generalized Maximum Likelihood	240
9.8.1	Definition of the Problem	240
9.8.2	Application of the Generalized Maximum Likelihood Criterion	241
9.8.3	Application of the Generalized Bayesian Rule	244
9.9	A Pattern Recognition Problem	245
9.9.1	An Illustrative Example	245

9.9.2	Application of the Sharp Decision Rule	246
9.9.3	Application of the Soft Decision Rule	246
9.10	Concluding Remarks.....	247
10.	Concluding Remarks and Outlook	249
	References	251
	Subject Index	255

